Does Experiential Probability Training Improve Numerical Decision Making?

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation
with honors research distinction in Psychology in the undergraduate colleges
of The Ohio State University

by

Kelsey Lewis

The Ohio State University
2014

Project Advisor: Dr. Ellen Peters, Department of Psychology

Abstract

Numeracy, the ability to understand and use mathematical concepts, is a crucial skill in which many lack proficiency. The less numerate are more likely to exhibit a number of cognitive biases that can result in suboptimal decision making. Two of these are probability insensitivity and attribute framing effects. Though prior research has shown a correlation between numeracy and the suppression of these phenomena, a causal link has yet to be determined. In an effort to demonstrate causation, the present study used a probability-training paradigm that targeted intuitive numeracy abilities. It was hypothesized that participants who received probability training would exhibit greater probability sensitivity and smaller attribute framing effects compared to untrained participants. Ohio State students ($N = 339$) were randomly assigned to one of two conditions. Participants in the training condition watched series of flashing faces, estimated the probabilities represented by each series, and received accurate feedback. Participants in the no training condition watched the same series but did not estimate probabilities nor receive probability-related feedback. All participants then judged the riskiness of a number of risk scenarios in which we manipulated the probability of an adverse outcome, allowing us to assess probability sensitivity. They also judged the attractiveness of several scenarios, either in the positive or negative frame, allowing us to assess framing effects. The experiment ended with a numeracy scale. Results reveal non-significant effects of experiential probability training on probability sensitivity and attribute framing effects overall, but effects of numeracy and training were found for some scenarios individually.

Keywords: numeracy, learning from experience, experiential probability training, attribute framing effects, probability sensitivity

Does Experiential Probability Training Improve Numerical Decision Making?

Although people make important decisions using quantitative information every day, many struggle to use numbers properly. Nineteen percent of U.S. adults cannot consistently and accurately carry out basic mathematical calculations involving counting, sorting, basic arithmetic, and simple percentages, such as 50% (OECD, 2013). This statistic highlights the prevalent deficit in numeracy, the ability to understand and use probabilistic and mathematical concepts (Peters, 2012). Deficits in numeracy become more alarming in light of their documented effects. On average, the less numerate are less healthy and less wealthy than their highly numerate peers (Gaglio, Glasgow, & Bull, 2012; McIntosh & Vignoles, 2001). Additionally, less numerate people's numerical decision making tends to be biased in predictable ways (Peters, 2012).

People generally show insufficient sensitivity to probabilities, especially when responding to nonmonetary scenarios (e.g., a health risk; McGraw, Shafir, & Todorov, 2010). We refer to this phenomenon as probability insensitivity. Less numerate people tend to be less sensitive to probabilities than highly numerate people. In one study investigating this difference, cancer patients were given a medical decision aid that predicted their cancer-free survival rates in the form of probabilities. Then, patients estimated their perceived cancer-free survival rates. In this study, highly numerate patients' perception of cancer-free survival was positively correlated with the decision aid's estimates, meaning that as the decision aid's probability of cancer-free survival increased, the patients' personal probability estimates also increased. However, less numerate patients' personal estimates remained constant regardless of the decision aid's increasing probabilities, demonstrating probability insensitivity (see Figure 1 in Appendix A; Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010). Because perceived survival rates influence patients' will-to-live, which has been shown to affect actual survival rates, probability insensitivity may have detrimental effects (Karppinen, Laakkonen, Strandberg, Tilvas, & Pitkälä, 2013).

Another way in which less numerate people's numerical decision making is biased involves attribute framing. Framing effects occur when logically equivalent outcomes are treated differently depending on the frame in which they are presented. Attribute framing involves a single attribute that is the subject of the framing (Levin, Schneider, & Gaeth, 1998). For example, attribute framing effects can be said to occur when a 75% lean burger is rated as better tasting and less greasy than a 25% fat burger, even though the information about the burgers is logically equivalent (Levin, Schneider, & Gaeth, 1998). The highly numerate are less susceptible to attribute framing, possibly because they are more likely to convert the given frames into alternative frames (i.e., the 75% into 25%) (Peters, 2012). The less numerate, however, are likely to change their judgments as the frame changes, rendering their decisions easier to manipulate by information providers.

As described, the less numerate are at a serious disadvantage compared to their highly numerate peers, especially because they are more susceptible to biases that negatively affect numerical decision making. Aids for the less numerate have been suggested, including the use of visual aids when presenting probabilistic information (Peters, 2012). However, visual aids facilitate decision making in temporary, highly specific contexts rather than improving decision making more generally. The present study investigates whether a numeracy intervention to teach probability understanding will improve numerical decision making across contexts, as demonstrated by the suppression of biases such as probability insensitivity and attribute framing effects. To date, only correlations between numeracy and probability sensitivity and numeracy and attribute framing effects have been documented (Peters, 2012). However, the present study aims to test a possible causal relationship such that improving numeracy will lead to greater probability sensitivity and smaller attribute framing effects.

Because less numerate people rely less heavily on numeric information to make decisions (Peters, 2012), we decided not to use explicit probability training. Instead, we use probability training that relies on a more intuitive, experiential learning process. Experiential learning of likelihoods occurs when the probability of an event is not given but instead must be inferred from repeated sampling of outcomes over time (Camilleri & Newell, 2013). People rely on this type of learning every day. For example, we are not told the exact probability of our car brakes working, but we have experientially learned that this probability is high (i.e., brakes will almost always stop the car).

We predict that experiential probability training will improve numerical decision making. Several experiments have provided evidence for this assertion. In one study, participants either experientially learned the probability of their children inheriting Down's syndrome or were explicitly given the probability. Participants who learned from experience rated themselves as less worried about the risk compared to participants who were explicitly given probabilities. The authors use this finding to suggest that experiential learning may improve the accuracy of probabilistic risk evaluations (Tyszka & Sawicki, 2011). A second study demonstrated that as probabilistic information becomes increasingly complex, participants prefer experiential learning to explicit learning, perhaps because learning from experience seems more intuitive (Lejarraga, 2010). A final study showed that training in an intuitive, approximate number system leads to improved accuracy in novel, concrete number manipulation tasks (Park & Brannon, 2013).

These studies suggest that people who undergo experiential probability training may exhibit improved numerical reasoning. If, as research suggests, improved numerical reasoning leads to improved decision making, experiential probability training should improve numerical decision making. We set out to test this. To conduct the training, we will use a method similar to that of Tyszka and Sawicki (2011) in which probabilities were represented by flashing pictures of children with or without Down's syndrome traits (e.g., upward slanted eyes and decreased muscle tone). However, our training will use more neutral pictures of yellow smiling and red frowning faces rather than pictures of children with or without Down's syndrome traits. Additionally, though participants in Tyszka and Sawicki's study (2011) experientially learned five probabilities, participants in our study will learn ten. These modifications were made because we want to teach a general understanding of probabilities, rather than an understanding specific to one context.

Our first hypothesis is that participants who undergo experiential probability training will exhibit greater probability sensitivity compared to participants who do not. Our second hypothesis is that participants who undergo experiential probability training will exhibit smaller attribute framing effects compared to participants who do not. We expect training to benefit all people but especially hope it to improve the numerical decision making of the less numerate. Significant results would provide evidence for causal relationships between numeracy and probability sensitivity and numeracy and attribute framing effects.

Method

Participants

Participants were 344 Ohio State undergraduate students who received academic credit for an introductory psychology course by participating in the study. Some participants' data (*n*=5) were excluded because they were not able to finish the survey due to computer issues, leaving 339 total participants. To the best of our knowledge, these data can be considered missing at random since none of the experimental variables correlated with the computer error. There were 161 males and 178 females, with a mean age of 18.65 years and an average of 12.47 completed years of education. 167 participants received the experiential probability training and 172 participants did not. Data were collected using Qualtrics, an online survey tool, and participants completed the survey in our laboratory, taking an average of 23 minutes.

Design

This study had three manipulated independent variables, including one training variable used in both experiments, one variable specific to the probability sensitivity experiment, and one specific to the framing experiment. This study also had two dependent variables, one in each experiment, and one individual difference measure, numeracy, which was used in both experiments. The first independent variable was whether or not participants received experiential probability training. Participants were randomly assigned to one of two conditions, making the training condition variable a between-subjects manipulation; about half of the participants were trained and about half were not.

The second independent variable was probability magnitude, which was used to assess probability sensitivity and was manipulated within subjects as a repeated measure. Five probability levels were embedded in and counterbalanced across five risk scenarios using a Latin squares design. The scenarios were presented to participants in random order. Each participant therefore saw all five scenarios and all five probability levels but did not see any scenarios or any probability levels more than once. The dependent variable for the probability sensitivity experiment was the reported amount of effort participants were willing to put forth to avoid the adverse outcome in each risk scenario (referred to as reported effort to avoid risk from here on).

The third independent variable was attribute frame. Frame had two levels, positive and negative, and was used to assess attribute framing effects. Frame was manipulated between-subjects, meaning that participants never saw two frames of the same scenario. However, each participant saw a mixture of positively and negatively framed scenarios. The dependent variable for the attribute framing experiment was quality judgments of certain characteristics described in the framing scenarios.

Procedure

After consenting to participate, participants were randomly assigned to the training condition or the no training condition. Participants in both conditions watched 11 series of flashing yellow smiling faces and red frowning faces. Each series represented 1 of 11 probabilities of a red frowning face. All participants first saw an example series representing a 70% probability of a red frowning face, and the response to this example series was not included in analysis. Then, participants saw the 10 experimental series of pictures, in random order, representing probabilities 2%, 7%, 11%, 18%, 22%, 27%, 32%, 35%, 41%, and 48%. Probabilities were chosen to reflect a representative, yet semi-random, sampling of probabilities ranging from 1% to 50%. 50% was chosen as a cutoff because larger probabilities have complementary probabilities below 50%. So, in a sense, the larger probabilities are represented by their smaller complements. The 11 series took around 12 minutes to complete.

Within each series, each smiling and frowning face appeared on the computer screen for 1 second. The total number of faces for each series, and thus, the duration of each series, varied. For example, one series contained 50 pictures and lasted 50 seconds while another contained 28 pictures and lasted 28 seconds. The total number of pictures varied to discourage participants in the training condition from calculating probabilities and to instead encourage intuitive estimation of probabilities. Additionally, the presentation order of the smiling and frowning faces in each series was randomized by the JavaScript used in Qualtrics.

The 167 participants randomly assigned to the training condition were instructed to watch each series and estimate the probabilities represented by the red frowning faces. The 172 participants randomly assigned to the no training condition were told nothing about probabilities. Instead, they were told to left click their mouse for every fifth picture that they saw, regardless of what kind of face it was. In this way, those in the training condition focused on estimating relative frequencies, in terms of probabilities and percentages, whereas those in the no training condition focused merely on count frequency.

Participants in the training condition responded with their probability estimate for each series in a free response box. Participants in the no training condition responded with the number of times that they should have left-clicked their mouse for each series, also in a free response box. After responding, all participants were given the correct answer and reminded of their response for comparison. If their response was correct, participants were congratulated and told that they earned 5 points. For the training condition, probability estimates were considered correct if they were within 5 percentage points, in either direction, from the exact answer. For example, a response to the 11% series was counted as correct if it was between 6% and 16%.

In contrast, the no training click response needed to be the exact answer to be considered correct because this task was much easier. If a response was incorrect, participants were told that they earned 0 points. They were informed that the points did not affect their academic credit for the study, but we hoped the points would motivate participants to attend to the series. Given that a few participants asked the experimenter for their point totals, the points seemed to motivate some participants at least to a degree.

Participants were then randomly assigned to one of five risk scenario and probability level combinations. Each combination contained the same five scenarios and the same five probability levels associated with the adverse outcomes described in the scenarios (1%, 9%, 17%, 23%, or 34%). Scenario topics included being robbed at gunpoint, caught in cold, heavy rain, falsely accused of cheating on an exam, involved in a car accident, and stuck with a slow-functioning laptop. The pairings of scenarios to probability levels were counterbalanced using a Latin square design so that each scenario was associated with each probability level among all five risk scenario combinations. This counterbalancing eliminated the effects of the severity of the adverse outcome described in the scenario (e.g., a car accident is more severe than getting caught in cold rain, regardless of the probability). As such, participants' judgments of the isolated probability levels and probability sensitivity were investigated. After the combination of scenarios and probability levels was chosen, the order of presentation was randomized. An example risk scenario is as follows: "While driving home in a residential neighborhood, a newspaper blows on your windshield. In this circumstance, 17% of people like you will get into a car accident." The remaining scenarios can be found in Appendix B.

After each risk scenario, participants responded to the first dependent measure: reported effort to avoid risk. They were asked, "In this scenario, how much effort would you put forth to guarantee that there is no chance of you _____?" For each scenario, the adverse outcome was described in the blank (e.g., "no chance of you getting into a car accident?"). The 5-point scale ranged from No Effort=0 to Extreme Effort=4.

Participants then responded to five framing scenarios in random order. Each scenario described some quality of a person or thing as a percentage. An example framing scenario is as follows: "In an Algebra course with 43 people, Peter took an online assessment that covered single variable equations. Peter answered 28% [72%] of the questions incorrectly [correctly]." In this example, 28% and the word "incorrectly" would appear in the negative frame of the scenario; 72% and the word "correctly" would appear in the positive frame of the scenario. For each scenario, the item qualities were associated with percentages (e.g., 6%, 8%, 12%, 21%, 28% or their complements 94%, 92%, 88%, 79%, 72%). The frame (positive or negative) determined which percentage was associated with the item qualities.

The frames were manipulated between subjects within each condition so that some participants saw the positive frame and some saw the negative of each scenario. Each participant saw a mixture of positive frames and negative frames: either three positive frames and two negative or three negative frames and two positive. Topics of the scenarios included a prescription drug with a probability of side effects or no side effects, an academic assessment with a percentage of answers incorrect or correct, milk that had some percent of fat content or some percent fat-free, health bars that had some percent of sugar content or some percent sugar-free, and donuts that had some percent of sugar content and some percent sugar-free. See Appendix C for the full text of each scenario.

After each framing scenario, participants rated the quality of something in the scenario, referred to as quality judgments. The quality that they rated was always relevant to the characteristic described with a percentage in the scenario. For example, in the scenario with Peter and the online assessment, participants rated Peter's performance on a 7-point scale from Extremely Poor=-3 to Extremely Good=3. The words of the scale varied for each scenario to encompass the quality described but consistently ranged from the most negative evaluation to the most positive.

Lastly, participants filled out the 8-item numeracy scale (Weller et al., 2012). This scale contains items testing probability in several formats (percentage, decimal, and ratio), Bayesian reasoning, Cognitive Reflection (Frederick, 2005), and other mathematical concepts. The scale was developed to have excellent psychometric properties based on a Rasch analysis and good predictive validity relative to existing scales. The scale can be found in Appendix D. Additionally, participants responded to four frame conversion items (see Appendix E), in random order, that corresponded to four of the framing scenarios. For these items, participants responded with the probability that was complementary to the one that they saw earlier in the framing scenarios. For example: “Peter answered 72% of his test questions correctly. What percentage of questions did he answer incorrectly?” Participants responded by typing a number into a free response box. These additional items assessed ability to convert frames in an effort to understand if frame conversion ability mediated numeracy’s mitigating effects on framing effects. Finally, participants answered demographic questions related to their gender, age, years of education, native language, GPA, ACT scores, and SAT scores.

Analysis

As previously mentioned, condition was either no training (coded as -0.5) or training (coded as 0.5). Two types of performance scores in both conditions were calculated. The first one, referred to as the mean difference score, was calculated by summing the absolute value differences between the correct answer for each series and each participant’s estimate for that series, then by taking the mean of those differences to create one number that reflected performance on all 10 series. The second performance score, referred to as the total correct, was a count of the number of correct responses across the 10 scenarios. As previously mentioned, participants' responses in the training condition were counted as correct if they were within 5 percentage points, in either direction, of the exact probability, whereas those in the no training condition needed to be the exact click count to be counted as correct.

In the risk scenarios, average reported effort to avoid risk was coded 0 to 4 with 0 meaning no effort would be put forth to avoid the adverse outcome. In the framing scenarios, quality judgments were coded -3 to 3 with the negative judgments reflecting low quality and positive judgments reflecting high quality. Frame was either negative (coded as -0.5) or positive (coded as 0.5). Frame conversion items were scored using the count of items answered correctly, with a 4 indicating that the participant answered all four frame conversion items correctly and a 0 indicating none (M = 3.76, Mdn = 4, SD = .60). Numeracy was scored using the count of items answered correctly, with an 8 indicating that the participant answered all eight items of the numeracy scale correctly and a 0 indicating none (M = 4.72, Mdn = 5, SD = 1.69). To present data comparing individuals lower and higher in numeracy, we used a median split (numeracy$\leq$5, numeracy$\geq$6). However, in the inferential analyses, numeracy was treated as a continuous variable.

Results

Experiential Probability Training

Performance scores in the training and no training conditions provide useful descriptive statistics. For the total correct, participants in the training condition provided the correct responses for, on average, 5.79 series out of 10 (SD = 1.69). For the mean difference score, participants in the training condition were, on average, 8.23 percentage points from the correct percentage (SD = 4.94). Within the training condition, less numerate participants were, on average, 8.80 percentage points from the correct percentage (*SD* = 5.05). More numerate participants were, on average, 7.18 percentage points from the correct percentage (*SD* = 4.59). A correlation between the mean difference scores in the training condition and numeracy was significant ($r(165) = -.22$, $p = .004$).

For the total correct, participants in the no training condition provided the correct responses for, on average, 7.94 series out of 10 (*SD* = 2.42). For the mean difference score, participants in the no training condition were, on average, .40 clicks from the correct click count (*SD* = .90). Because most participants did well on this easier task, comparing the mean difference score to numeracy is perhaps not as meaningful as comparing the total correct to numeracy. Within the no training condition, less numerate participants provided the correct responses for, on average, 7.62 series out of 10 (*SD* = 2.55). More numerate participants provided the correct responses for, on average, 8.60 series out of 10 (*SD* = 2.00). An independent samples t-test was conducted to compare the total correct for less numerate and more numerate participants in the no training condition, and it was significant ($t(170) = 2.54$, $p = .01$).

Probability Sensitivity

We conducted a series of mixed model regressions assessing reported effort to avoid risk among the five risk scenarios simultaneously. These models allow observations to correlate both within subject and within scenario. The first model, with reported effort to avoid risk as the dependent variable (DV) and probability level as the independent variable (IV), revealed that participants increased their reported effort to avoid risk as the probability levels of the adverse outcomes increased; this main effect was significant ($t = 13.86$, $p < .001$). In a second model, we tested for order effects for the risk scenarios, even though the scenarios were randomly presented to participants. With reported effort to avoid risk as the DV and the interaction between order of scenario presentation and probability level and their main effects as IVs, the interaction was not significant. Therefore we removed order from further models and do not discuss order effects further (t = -.81, p = .42).

A third model, with reported effort to avoid risk as the DV, included the interaction between probability level and numeracy as well as their main effects as IVs. We found a significant interaction such that the highly numerate were more sensitive to probabilities than the less numerate across the five risk scenarios (t = 2.02, p = .04), replicating previous results using several new scenarios (Peters, 2012). The interaction seems to occur because the less numerate perceive more risk than the highly numerate at the lower probabilities. See Table 1 for a summary of the numeracy by probability level interaction. See Figure 2 for a visual depiction of these results. When the analysis for the third model was conducted separately for each scenario using linear regressions, only one scenario about being falsely accused of cheating returned a marginally significant interaction between probability level and numeracy (t = 1.66, p = .098). However, three of the other four scenarios showed interactions in the right direction, explaining why the overall model was significant. See Table 2 for a summary of the scenario-by-scenario probability level by numeracy interaction.

Table 1

Mixed Model Regression Three Using Reported Effort to Avoid Risk as the Dependent Variable and Main Effects of Probability Level, Numeracy, and Their Interaction as Independent Variables

	B	*t*	*p*
Probability Level	.016	2.77	= .01
Numeracy	-.090	-2.94	= .003
Probability X Numeracy	.002	2.02	= .04

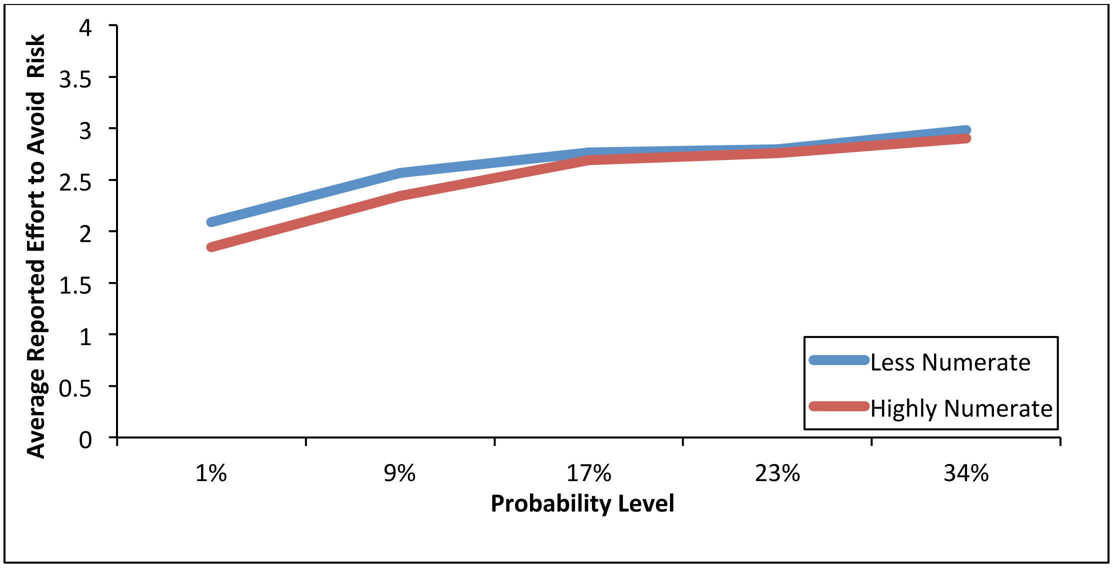

Figure 2. Average reported effort to avoid risk by numeracy and probability level across the five scenarios.

Table 2

Scenario-by-Scenario Linear Regression Analyses Using Reported Effort to Avoid Risk as the Dependent Variable and Main Effects of Probability Level, Numeracy, and Their Interaction as Independent Variables

	B	*t*	*p*
Scenario 1: Robbed at gunpoint			
Probability Level	-.004	-.22	= .83
Numeracy	-.141	-2.04	= .04
Probability X Numeracy	.005	1.55	= .12
Scenario 2: Caught in cold, heavy rain			
Probability Level	.015	.95	= .34
Numeracy	-.166	-2.76	= .01
Probability X Numeracy	.002	.66	= .51
Scenario 3: Falsely accused of cheating on an exam			
Probability Level	.004	.28	= .78
Numeracy	-.104	-1.69	= .09
Probability X Numeracy	.005	1.66	= .098
Scenario 4: Car accident			
Probability Level	.016	.99	= .32
Numeracy	-.098	-1.51	= .13
Probability X Numeracy	.003	.92	= .36
Scenario 5: A slow-functioning laptop			
Probability Level	.053	3.56	< .001
Numeracy	.069	1.15	= .25
Probability X Numeracy	-.004	-1.41	= .16

Note. Scenario 1: Adjusted $R^2 = .05$, $F = 6.5$, $p < .001$. Scenario 2: Adjusted $R^2 = .09$, $F = 12.7$, $p < .001$. Scenario 3: Adjusted $R^2 = .08$, $F = 10.5$ $p < .001$. Scenario 4: Adjusted $R^2 = .08$, $F = 10.7$, $p < .001$. Scenario 5: Adjusted $R^2 = .11$, $F = 14.9$, $p < .001$.

According to Hypothesis 1, participants in the training condition should exhibit greater probability sensitivity compared to participants in the no training condition. In a fourth model, we therefore predicted reported effort to avoid risk (DV) with the interaction between probability level and training condition, as well as their main effects. The interaction was not significant ($t = 1.27$, $p = .20$). In other words, trained and untrained participants' overall reported effort to avoid risk was similarly influenced by the changing probability levels. See Table 3 for a summary and Figure 3 in Appendix A for a visual representation of the fourth mixed model regression results.

Table 3

Mixed Model Regression Four Using Reported Effort to Avoid Risk as the Dependent Variable and Main Effects of Probability Level, Training Condition, and Their Interaction as Independent Variables

	B	*t*	*p*
Probability Level	.027	13.88	< .001
Training Condition	-.236	-2.28	= .02
Probability X Training	.005	1.27	= .20

In a series of linear regressions, we conducted the analysis in the fourth model separately for each scenario. The results, detailed in Table 4, indicated that one of the scenarios about being caught in cold, heavy rain did have a strong significant interaction between probability level and training condition in the predicted direction ($t = 2.88$, $p = .004$). The other scenarios showed no significant interactions. See Figure 4 for a visual representation of the significant interaction for the rain scenario.

Table 4

Scenario-by-Scenario Linear Regression Analyses Using Reported Effort to Avoid Risk as the Dependent Variable and Main Effects of Probability Level, Training Condition, and Their Interaction as Independent Variables

	B	*t*	*p*
Scenario 1: Robbed at gunpoint			
Probability Level	.020	3.81	< .001
Training Condition	-.251	-1.16	= .25
Probability X Training	-.003	-.27	= .79
Scenario 2: Caught in cold, heavy rain			
Probability Level	.024	4.66	< .001
Training Condition	-.524	-2.49	= .01
Probability X Training	.030	2.88	= .004
Scenario 3: Falsely accused of cheating on an exam			
Probability Level	.027	5.26	< .001
Training Condition	-.066	-.32	= .75
Probability X Training	-.005	-.46	= .65
Scenario 4: Car accident			
Probability Level	.0297	5.39	< .001
Training Condition	-.0287	-.13	= .90
Probability X Training	-.000	-.00	= .999
Scenario 5: A slow-functioning laptop			
Probability Level	.033	6.59	< .001
Training Condition	-.339	-1.65	= .01
Probability X Training	.004	-.41	= .68

Note. Scenario 1: Adjusted $R^2 = .05$, $F = 7.2$, $p < .001$. Scenario 2: Adjusted $R^2 = .08$, $F = 10.3$, $p < .001$. Scenario 3: Adjusted $R^2 = .07$, $F = 10.0$, $p < .001$. Scenario 4: Adjusted $R^2 = .07$, $F = 9.7$, $p < .001$. Scenario 5: Adjusted $R^2 = .12$, $F = 16.3$, $p < .001$.

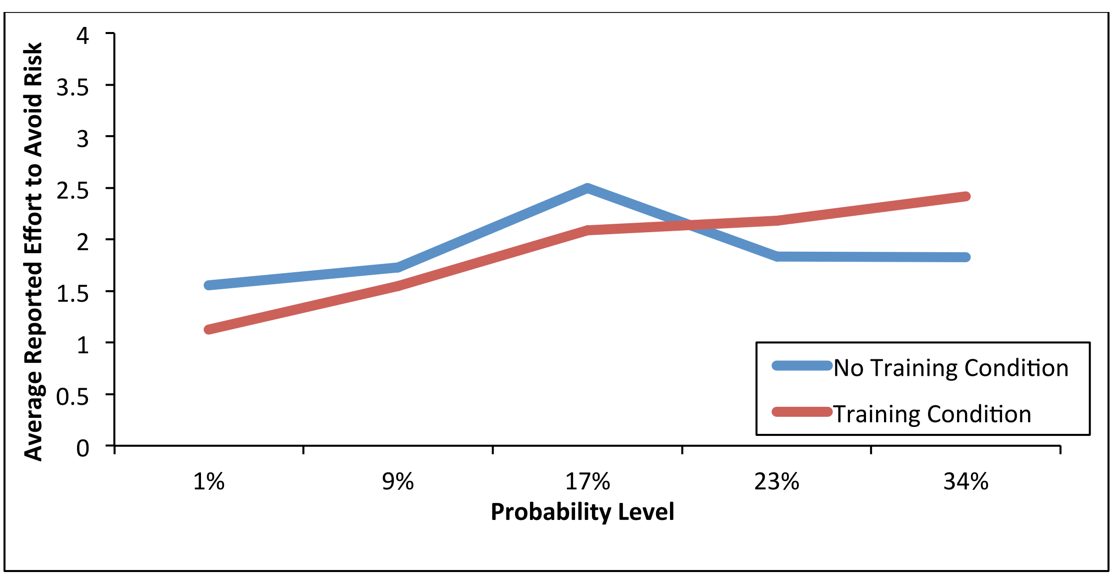

Figure 4. Average reported effort to avoid risk by training condition and probability level for the rain scenario.

Though we did not predict a three-way interaction between probability level, training condition, and numeracy, we conducted a fifth mixed model analysis to determine if significant differences existed since differences in aptitude for numeric processing may have altered the effects of training. The three-way interaction, with reported effort to avoid risk as the DV and the main effects of probability level, training condition, numeracy, and their interactions as IV's, was non-significant ($t = 0.54$, $p = .59$). See Table 5 for a summary of this non-significant three-way interaction.

Table 5

Mixed Model Regression Five Using Reported Effort to Avoid Risk as the Dependent Variable and Main Effects of Probability Level, Training Condition, Numeracy, and Their Interactions as Independent Variables

	B	*t*	*p*
Probability Level	.016	2.70	= .01
Training Condition	-.358	-1.16	= .25
Numeracy	-.087	-2.84	= .01
Probability X Training	-.001	-.10	= .92
Probability X Numeracy	.002	2.05	= .04
Training X Numeracy	.027	.43	= .66
Probability X Train X Num	.001	.54	= .59

Between-Subjects Attribute Framing Effects

A mixed model regression with quality judgments as the dependent variable and frame as the independent variable revealed that participants judged positively framed scenarios ($M = 0.53$) to be superior to negatively framed scenarios ($M = -0.45$), and that this difference was significant ($t = 13.49$, $p < .001$). Since all participants saw a mixture of positive and negative frames, one could argue that seeing both may have caused participants to think more about complementary frames. Thus, the framing effects may have exhibited order effects. However, a second mixed model regression with quality judgments as the dependent variable and the interaction between frame and the order of presentation and their main effects as independent variables revealed an overall non-significant interaction between quality judgments and the order of framing scenarios (t = -1.06, p = .29). We therefore did not include order of presentation in subsequent analyses.

In a third mixed model regression with the interaction of numeracy and frame and their main effects as independent variables, we found no significant interaction, failing to replicate previous results on new scenarios (Peters et al., 2006; see Table 6). The differences between less numerate people's mean judgments of -.39 and .61 and the more numerate people's mean judgments of -.56 and .38 in the negative and positive frame respectively were not significant, as indicated by the interaction of numeracy and frame (t = -.79, p = .43).

Table 6

Mixed Model Regression Three Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Numeracy, and Their Interaction as Independent Variables

	B	*t*	*p*
Frame	1.124	5.29	< .001
Numeracy	-.063	-2.39	= .02
Frame X Numeracy	-.034	-.79	= .43

Such null results have previously been reported by Sinayev (2013), but he did replicate the numeracy by frame interaction for the scenario used by Peters et al. (2006). To check for replication in our data, we conducted linear regression analyses of the five scenarios separately, revealing a significant interaction between frame and numeracy for the academic assessment scenario (see Appendix C for the full text of this scenario). In the academic assessment scenario, the highly numerate exhibited smaller framing effects as expected and as found by Sinayev

(2013) and Peters et al. (2006). The less numerate made mean judgments of -.54 and .04, and the highly numerate made mean judgments of -.42 and -.35 in the negative and positive frame respectively, with a significant interaction of numeracy and frame (t = -2.00, p = .05). See Table 7 for a summary of the numeracy by frame interactions for each scenario. See Figure 5 for a visual representation of the numeracy by frame significant interaction for the academic assessment scenario.

Additionally, there was a marginally significant interaction in the opposite direction expected for the donut scenario so that the less numerate exhibited smaller framing effects than the highly numerate. For this scenario, the less numerate made mean judgments of -.88 and -.07, and the highly numerate made mean judgments of -1.07 and .34 in the negative and positive frame respectively, with a marginally significant interaction of numeracy and frame (t = 1.73, p = .09).

Table 7

Scenario-by-Scenario Linear Regression Analyses Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Numeracy, and Their Interaction as Independent Variables

	B	*t*	*p*
	Scenario 1: Prescription drug		
Frame	.946	2.15	= .03
Numeracy	-.088	-2.00	= .05
Frame X Numeracy	.042	.48	= .64
	Scenario 2: Academic assessment		
Frame	1.314	2.70	= .01
Numeracy	-.051	-1.04	= .30
Frame X Numeracy	-.194	-2.0	= .05
	Scenario 3: Milk		
Frame	1.205	2.31	= .02
Numeracy	-.097	-1.86	= .06
Frame X Numeracy	-.008	-.08	= .94
	Scenario 4: Health bars		
Frame	1.915	3.79	< .001

Numeracy	-.140	-2.78	= .01
Frame X Numeracy	-.161	-1.60	= .11
		Scenario 5: Donuts	
Frame	.153	.29	= .77
Numeracy	.026	.50	= .62
Frame X Numeracy	.181	1.73	= .09

Note. Scenario 1: Adjusted $R^2 = .15$, $F = 20.9$, $p < .001$. Scenario 2: Adjusted $R^2 = .02$, $F = 3.7$, $p = .01$. Scenario 3: Adjusted $R^2 = .12$, $F = 15.7$, $p < .001$. Scenario 4: Adjusted $R^2 = .14$, $F = 19.2$, $p < .001$. Scenario 5: Adjusted $R^2 = .09$, $F = 12.1$, $p < .001$.

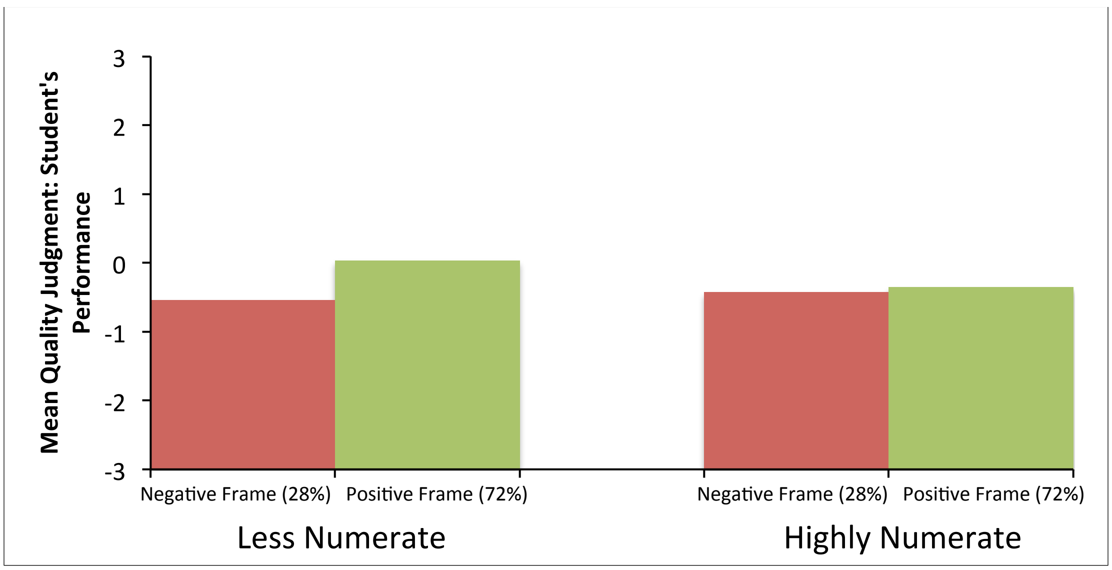

Figure 5. Average quality judgments by numeracy and frame for the academic assessment scenario.

According to Hypothesis 2, framing effects should be smaller for the training condition than the no training condition. In a fourth mixed model regression, with quality judgments as the dependent variable and main effects of frame, training condition, and their interaction as IV's, the interaction term was not significant. Participants in the training condition made mean judgments of -.34 and .58, and participants in the no training condition made mean judgments of -.55 and .48 in the negative and positive frame respectively, with the interaction of training condition and frame not approaching significance ($t = -.35$, $p = .72$). See Table 8 for a summary of the overall frame by training interaction.

Table 8

Mixed Model Regression Four Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Training Condition, and Their Interactions as Independent Variables

	B	*t*	*p*
Frame	.965	13.49	< .001
Training Condition	.151	1.69	= .09
Frame X Training	-.051	-.35	= .72

Furthermore when the interaction between frame and training was assessed in each scenario individually, using a series of linear regressions, there were no significant effects. However, the academic assessment scenario did show a directional effect (interaction $p = .10$). See Table 9 for a summary of the training by frame interactions for each scenario and Figure 6 for a visual representation of the frame by training interaction for the academic assessment scenario.

Table 9

Scenario-by-Scenario Linear Regression Analyses Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Training Condition, and Their Interaction as Independent Variables

	B	*t*	*p*
	Scenario 1: Prescription drug		
Frame	1.132	7.60	< .001
Training Condition	.032	.22	= .83
Frame X Training	-.302	-1.01	= .31
	Scenario 2: Academic assessment		
Frame	.395	2.41	= .02
Training Condition	.0878	.54	= .59
Frame X Training	-.535	-1.63	= .10
	Scenario 3: Milk		
Frame	1.18	6.71	< .001
Training Condition	.29	1.67	= .095
Frame X Training	.29	.82	= .42

		Scenario 4: Health bars	
Frame	1.164	6.78	< .001
Training Condition	.170	.99	= .32
Frame X Training	.407	1.19	= .24
		Scenario 5: Donuts	
Frame	1.026	5.79	< .001
Training Condition	.228	1.29	= .20
Frame X Training	-.101	-.29	= .78

Note. Scenario 1: Adjusted $R^2 = .14$, $F = 19.7$, $p < .001$. Scenario 2: Adjusted $R^2 = .02$, $F = 3.0$, $p = .03$. Scenario 3: Adjusted $R^2 = .12$, $F = 15.7$ $p < .001$. Scenario 4: Adjusted $R^2 = .12$, $F = 16.4$, $p < .001$. Scenario 5: Adjusted $R^2 = .09$, $F = 11.5$, $p < .001$.

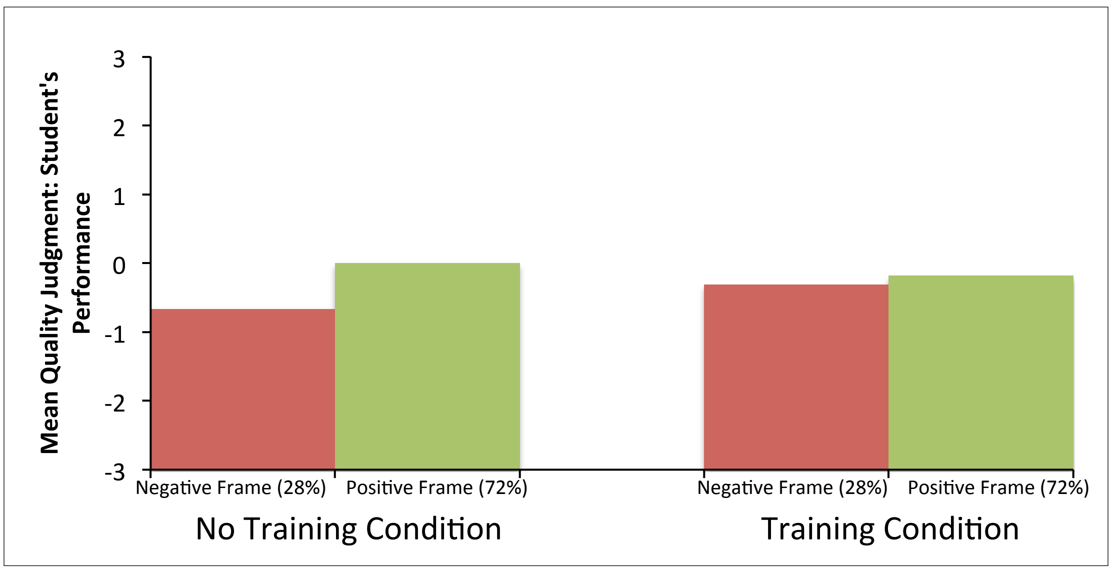

Figure 6. Average quality judgments by training condition and frame for the academic assessment scenario.

Once again, although we did not predict a three-way interaction between frame, training condition, and numeracy, we conducted a fifth mixed model analysis to determine if aptitude for numeric processing altered the effects of training. The three-way interaction, with quality judgments as the DV and main effects of frame, training condition, numeracy, and their interactions as IV's, was not significant ($t = .52$, $p = .61$). See Table 10 for a summary of this non-significant three-way interaction.

Table 10

Mixed Model Regression Five Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Numeracy, Training Condition, and Their Interactions as Independent Variables

	B	*t*	*p*
Frame	1.107	5.16	$< .001$
Training Condition	.0572	.21	$= .83$
Numeracy	-.063	-2.36	$= .02$
Frame X Training	-.262	-.61	$= .54$
Frame X Numeracy	-.030	.70	$= .49$
Training X Numeracy	.020	.38	$= .70$
Frame X Train X Num	.044	.52	$= .61$

Frame Conversion Items

In a sixth model using a linear regression analysis, with frame conversion scores as the dependent variable and training condition as the independent variable, the main effect was not significant ($t = -.66$, $p = .51$). In other words, trained participants frame conversion scores ($M = 3.75$) did not significantly differ from untrained participants frame conversion scores ($M = 3.77$). In a seventh model using a linear regression analysis, with frame conversion scores as the dependent variable and numeracy as the independent variable, the main effect was significant (t = 12.11, p < .001). Highly numerate participants' frame conversion scores ($M = 3.91$) were significantly higher than less numerate participants' frame conversion scores ($M = 3.68$).

We also looked at the frame conversion items to determine if framing effects were influenced by people's ability to convert given frames into complement frames. Because most participants answered all four frame conversion items correctly (82% or $n = 278$; a possible ceiling effect), we did not expect a significant effect, but conducted the analysis regardless. In an eighth mixed model regression, with quality judgments as the dependent variable and the interaction between frame conversion scores and frame and their main effects as independent variables, the interaction term was not significant (t = -.52, p = .60). We ran linear regression analyses of the five scenarios separately testing for the same interaction between frame conversion scores and frame. We found a marginally significant interaction for the academic assessment scenario (t = -1.76, p = .08). See Table 11 in Appendix A. We then conducted a Sobel Mediation Test to see if the frame conversion ability mediated the interaction between numeracy and frame for the academic assessment scenario. The mediation test did not approach significance (z = .50, p = .62).

Discussion

This study investigated a causal link between training intuitive numeracy skills and the suppression of numerical decision-making biases, such as probability insensitivity and attribute framing effects. It was hypothesized that participants who received experiential probability training, a training that targeted intuitive probabilistic understanding, would exhibit greater probability sensitivity (Hypothesis 1) and smaller framing effects (Hypothesis 2).

Contrary to Hypothesis 1, trained participants were not more sensitive to probabilities on average than untrained participants across the risk scenarios. However, consistent with Hypothesis 1, trained participants were significantly more sensitive to probabilities in the rain scenario specifically. In other words, for this scenario, they demonstrated greater differences in the average reported effort they would put forth to avoid the risk as the probability of the risk increased compared to untrained participants.

Although the expected interaction was absent overall, we did find a main effect such that trained participants reported lower effort to avoid risk than the untrained participants; this finding may be consistent with that of Tyszka and Sawicki (2011) who found that participants who experientially learned probabilities were less worried about the risk of Down’s Syndrome compared to those who explicitly learned probabilities. So, it seems as though the experiential probability training reduced risk perceptions across the five probability levels (see Figure 3 in Appendix A).

Although there is no optimal level of risk perception, training's reduction in risk perceptions seem to match the effects of numeracy on risk perceptions. For numeracy, a similar main effect was found where the highly numerate participants reported lower effort to avoid risk than the less numerate participants. So, in our study, the average reported effort to avoid risk was lower and very similar for the trained and highly numerate, and higher and very similar for the untrained and the less numerate (see Figure 7 in Appendix A). These results are suggestive of the experiential probability training helping to close the gap between the risk judgments of the less numerate and highly numerate. These results also suggest that lower risk perceptions may correlate with improved understanding of numbers, perhaps implying some benefit of lower risk perceptions for probability sensitivity, such as more extensive focus on the probabilities themselves. Future research should examine whether more expansive training that includes additional probabilities leads to even lower risk perceptions and even greater probability sensitivity.

Contrary to Hypothesis 2, trained participants did not exhibit smaller framing effects on average than untrained participants. However, consistent with Hypothesis 2, the academic assessment scenario revealed directional effects such that trained participants exhibited smaller framing effects. In other words, the differences in trained participants' average quality judgments between the negative and positive frames were smaller compared to untrained participants' judgments. Taking the results of the two biases together, this study provides some evidence for the claim that improving numeracy, via experiential probability training, causes improved numerical decision making, as demonstrated by less expression of numerical decision-making biases in some scenarios. This possibility is further supported because the academic assessment scenario was the only one in which significant numeracy interactions with framing conditions emerged. Thus, only where a numeracy by frame interaction existed, training appeared to attenuate the bias.

It is possible that significant effects of training were not found for some of the scenarios because the no training condition contained a numerical component too similar to the training condition. Specifically, both conditions caused participants to think about and work with numbers. Perhaps getting participants to think about numbers was enough to improve their numerical decision making. Alternatively, perhaps getting participants to think about numbers improved their confidence in their number use, which then lead to improved numerical decision making. Because we did not include a control condition that did not contain a numerical component, we cannot be sure that our effects are due to numerical training and that our lack of effects is due to a lack of numerical training. Future studies should include the click count control condition and an additional control condition that completes a non-numerical task. Additionally, the subjective numeracy scale (Fagerlin et al., 2007), which assesses people's perceptions of their mathematical abilities, could be included as a sort of numerical confidence measure.

It is also possible that significant effects of training were not found in some risk and framing scenarios because of something unique to the scenarios. For the risk scenarios, training was found to significantly improve probability sensitivity only in the rain scenario. Moreover, this effect did not seem to be due to statistical randomness given its robustness. In a graph depicting the scenario-by-scenario sizes of interaction coefficients for the interaction between probability level and training condition, the rain scenario clearly appears to be affecting the interaction differently than the other scenarios (see Figure 8 in Appendix A). Because the highly numerate demonstrated greater average probability sensitivity across all five scenarios, the lack of effects for a training that targets numeracy abilities cannot be explained by a lack of numeracy effects.

For the rain scenario, the sample as a whole responded with less effort to avoid the adverse outcome compared to the other scenarios (see Figure 9 in Appendix A for a graph of scenario means). Therefore, perhaps trained participants were able to focus more on the probability levels in the rain scenario because the adverse outcome, i.e., getting drenched in cold, heavy rain, was seen as safer. If trained participants were too distracted by the severity of the adverse outcomes in the other scenarios to attend to the probability levels, we would expect to see no differences between conditions. Contrary to this line of reasoning, however, the laptop scenario, which also showed relatively small average reported effort to avoid the risk, did not display the interaction between training and risk probability even directionally.

For the framing scenarios, training was found to significantly lessen framing effects only in the academic assessment scenario. Because this is the only scenario in which significant effects of numeracy were found and the training targets numeracy abilities, this finding makes sense. However, why did numeracy affect framing effects only in the academic scenario in both our study and a previous study by Sinayev (2013)? We believe it has something to do with frame conversions, as theorized by Peters and Levin (2008), and motivation to do these conversions as outlined below.

In our study, we measured people's ability to perform frame conversions. In our regressions, it appeared that general numeracy accounts for any relationship between ability to do conversions and the size of the framing effect, rather than vice versa, since conversion ability was not a significant mediator of the effect. Our results, of course, do not rule out frame conversions as a mediator, but instead suggest that motivation to convert frames, rather than ability, might mediate this documented effect. In other words, at least in our college sample, attribute framing effects may be the result of not caring enough to make a conversion, rather than not being able to do so.

If motivation to convert frames does mediate numeracy's effect on attribute framing effects, then perhaps the highly numerate exhibited smaller framing effects only in the academic assessment scenario because they were motivated to convert frames only in this scenario. Because we know that relevance is positively correlated with motivation (Frymier & Shulman, 1995), increased motivation to convert frames may be found in the academic assessment scenario because academics are extremely relevant to students participating in research for class credit. In fact, students have to make this conversion often because some teachers return tests scored in terms of the number correct whereas other teachers score in terms of the number incorrect. In contrast, the other scenario subjects, prescription drugs, milk, health bars, and donuts, may have seemed significantly less relevant at the time; thus, even highly numerate participants were perhaps unmotivated to convert frames.

In conclusion, our study provides some supporting evidence for a causal link between intuitive numeracy skills and the suppression of numerical decision-making biases, such as probability insensitivity and attribute framing effects. It also provides a starting point for developing an experiential probability training paradigm that improves intuitive numeracy skills. However, additional studies need to be conducted to better understand the contexts in which the training works. Eventually, with additional evidence of its general efficacy, experiential probability training may be implemented to improve people's numerical decision making in various contexts, particularly in medical contexts where crucial life decisions involving numerical information are made.

References

Camilleri, A. R. & Newell, B. R. (2013). The long and short of it: Closing the description-experience "gap" by taking the long run view. *Cognition, 126*, 54-71.

Fagerlin A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: development of the subjective numeracy scale. *Medical Decision Making, 27*(5), 672–80.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

Frymier, A. B., & Shulman, G. G. (1995). How important is relevance to student motivation?. *Journal Of Adolescent & Adult Literacy*, *39*(3), 250.

Gaglio, B., Glasgow, R. E., & Bull, S. S. (2012). Do patient preferences for health information vary by health literacy or numeracy? A qualitative assessment. *Journal of Health Communication, 17,* 109-121.

Huizinga, M. M., Beech, B. M., Cavanaugh, K. L., Elasy, T. A., & Rothman, R. L. (2008). Low numeracy skills are associated with higher BMI. *Obesity, 16*, 1966-1968.

Karppinen, H., Laakkonen, M., Strandberg, T. E., Tilvis, R. S., & Pitkälä, K. H. (2012). Will-to-live and survival in a 10-year follow-up among older people. *Age & Ageing*, *41*, 789-794.

Lejarraga, T. (2010). When experience is better than description: Time delays and complexity. *Journal of Behavioral Decision Making, 23,* 100-116.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76*, 149-188.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Medical Decision Making, 30,* 464-473.

McGraw, A. P., Shafir, E., & Todorov, A. (2010). Valuing money and things: Why a $20 item can be worth more and less than $20. *Management Science, 56,* 816-830.

McIntosh, S. & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers, 53,* 453-481.

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing.

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*(10), 2013-2019.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21,* 31-35.

Peters, E., & Levin, I. P. (2008). Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options. *Judgment and Decision Making*, *3*(6), 435-448.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*, 408-414.

Sinayev, A. (2013). *Numeric conversions ameliorate the framing effect*. (Master's Thesis). Retrieved from https://etd.ohiolink.edu/

Tyszka, T., & Sawicki, P. (2011). Affective and cognitive factors influencing sensitivity to probabilistic information. *Risk Analysis, 31,* 1832-1845.

Weller, J., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W., & Peters, E. (2012).

Development and testing of an abbreviated numeracy scale: A rasch analysis approach. *Journal of Behavioral Decision Making*.

Appendix A

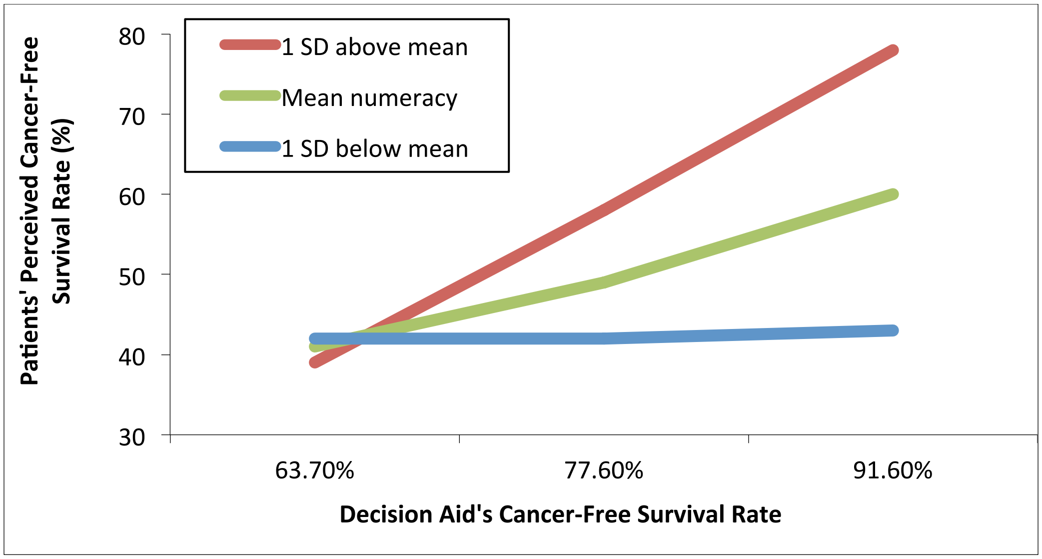

Figure 1. Average perceived cancer-free survival rates by numeracy and decision aid's cancer-free survival rates (Lipkus, Peters, Kimmick, Liotcheva, & Marcom, 2010).

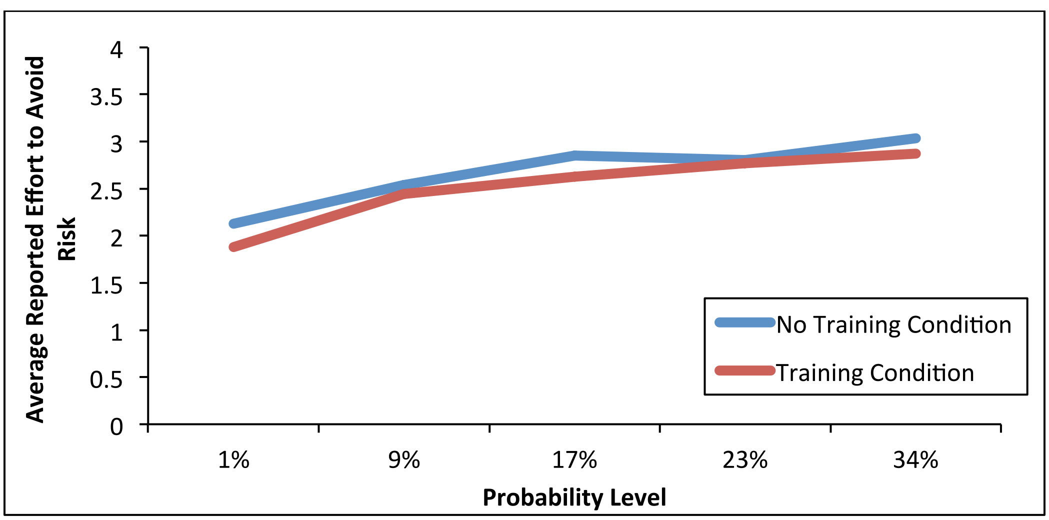

Figure 3. Average reported effort to avoid risk by training condition and probability level across the five scenarios.

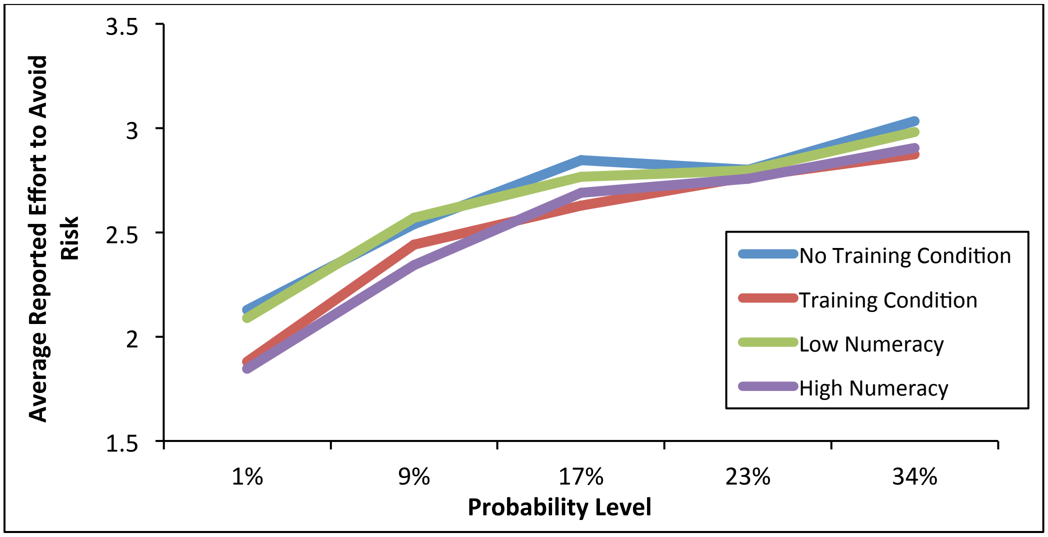

Figure 7. Average reported effort to avoid risk by probability level and differentiated by numeracy and training condition.

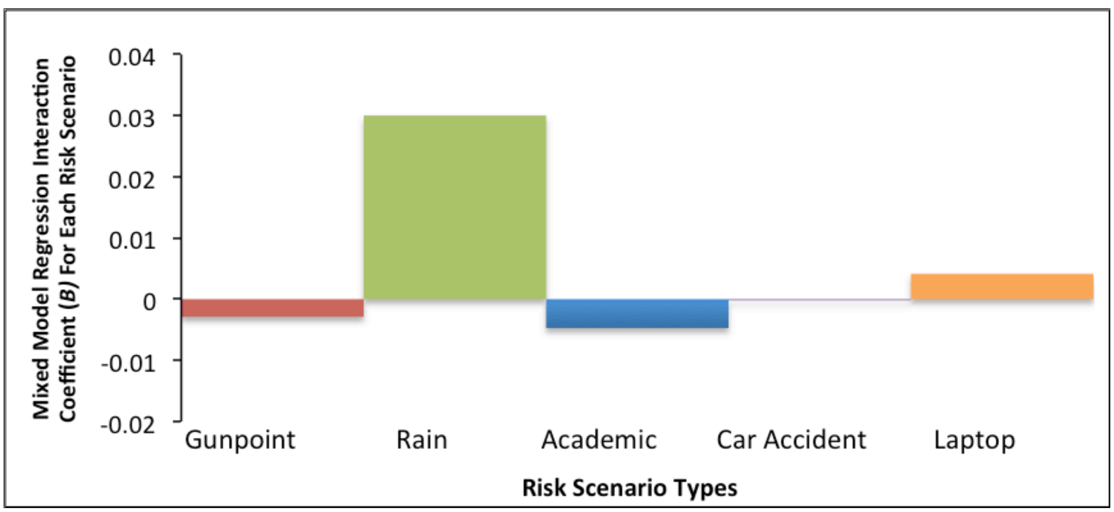

Figure 8. Scenario-by-scenario mixed model regression interaction coefficients for the interaction between probability level and training condition.

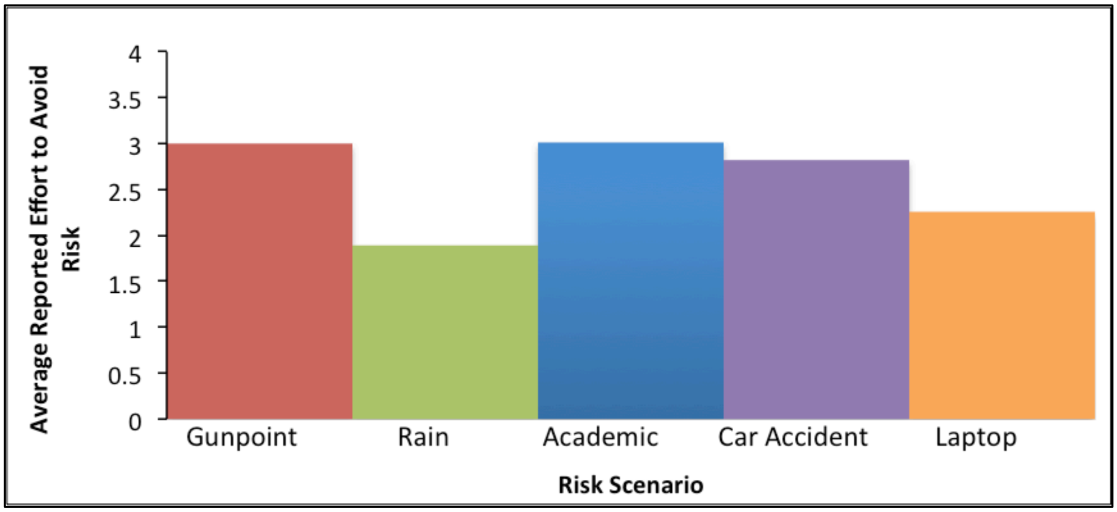

Figure 9. Scenario-by-scenario average reported effort to avoid risk.

Table 11

Scenario-by-Scenario Linear Regression Analyses Using Quality Judgments as the Dependent Variable and Main Effects of Frame, Frame Conversion Scores, and Their Interaction as Independent Variables

	B	*t*	*p*
	Scenario 1: Prescription drug		
Frame	1.080	1.10	= .27
Frame Conversion	.025	.20	= .84
Frame X Conversion	.014	.05	= .96
	Scenario 2: Academic assessment		
Frame	2.278	2.12	= .03
Frame Conversion	.039	.28	= .78
Frame X Conversion	-.498	-1.76	= .08
	Scenario 3: Milk		
Frame	1.710	1.48	= .14
Frame Conversion	.046	.30	= .76
Frame X Conversion	-.148	-.49	= .63
	Scenario 4: Health bars		
Frame	1.280	1.15	= .25
Frame Conversion	-.127	-.85	= .39
Frame X Conversion	-.033	-.11	= .91
	Scenario 5: Donuts		
Frame	.023	.02	= .98
Frame Conversion	.247	1.62	= .11
Frame X Conversion	.263	.86	= .39

Note. Scenario 1: Adjusted $R^2 = .14$, $F = 19.3$ $p < .001$. Scenario 2: Adjusted $R^2 = .02$, $F = 3.2$, $p = .02$. Scenario 3: Adjusted $R^2 = .11$, $F = 14.5$ $p < .001$. Scenario 4: Adjusted $R^2 = .12$, $F = 15.8$, $p < .001$. Scenario 5: Adjusted $R^2 = .09$, $F = 12.4$, $p < .001$.

Appendix B

Risk Scenarios

Scenario 1: Robbed at gunpoint

"You are walking through a relatively dangerous area after a party that ended late at night and see some suspicious characters. In this circumstance, _% of people like you will get robbed at gunpoint."

Scenario 2: Caught in cold, heavy rain

"The temperature is just below freezing and you left for class in a hurry, dressed in a sweater, without an umbrella. You have to walk a mile to get back home. In this circumstance, _% of people like you will be caught in heavy rain."

Scenario 3: Falsely accused of cheating on an exam

"Before taking an exam, you wrote some things that were unrelated to it on your hand, and your professor noticed this. In this circumstance, _% of people like you will be punished despite not having cheated."

Scenario 4: Car accident

"While driving home in a residential neighborhood, a newspaper blows on your windshield. In this circumstance, _% of people like you will get into a car accident."

Scenario 5: A slow-functioning laptop

"You are downloading new software for your new laptop, but it may be interfering with the hardware. In this circumstance, _% of people like you will end up with a laptop that works twice as slowly as it did before as a result of this interference."

Appendix C

Framing Scenarios

Scenario 1: Prescription drug

“Clearex is an acne cream available by prescription only in 12.8 ounce bottles. In clinical trials it caused [no] side effects in 8% [92%] of patients.”

Scenario 2: Academic assessment

“In an Algebra course with 43 people, Peter took an online assessment that covered single variable equations. Peter answered 28% [72%] of the questions incorrectly [correctly].”

Scenario 3: Milk

“White Valley is a new dairy company, founded in 2008. It aims to corner a low cost market by selling only whole milk, which is about 12% [88%] fat[-free].”

Scenario 4: Health bars

“Health Bars can be bought in bulk if at least 131 of them are purchased. Each bar is a combination of nuts and dark chocolate, made with pure cane sugar. Each bar is 21% [79%] sugar[-free].”

Scenario 5: Donuts

“Special Os are sold in paper cartons and can be bought in bulk if more than 259 are purchased. They are small donuts covered in cinnamon or vanilla, usually eaten with milk. Each pack is 6% [94%] sugar[-free].”

Appendix D

8-item Numeracy Scale

1. Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up as an even number?
2. 2. In the BIG BUCKS LOTTERY, the chances of winning a $10.00 prize are 1%. What is your best guess about how many people would win a $10.00 prize if 1,000 people each buy a single ticket from BIG BUCKS?
3. In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets of ACME PUBLISHING SWEEPSTAKES win a car?
4. If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1000?
5. If the chance of getting a disease is 20 out of 100, this would be the same as having a ____% chance of getting the disease.
6. Suppose your friend just had a mammogram. The doctor knows from previous studies that, of 100 women like her, 10 have tumors and 90 do not. Of the 10 who do have tumors, the mammogram correctly finds 9 with tumors and incorrectly says that 1 does not have a tumor. Of the 90 women without tumors, the mammogram correctly finds 80 without tumors and incorrectly says that 10 have tumors. The table below summarizes this information. Imagine that your friend tests positive (as if she had a tumor), what is the likelihood that she actually has a tumor?

	Tested Positive	Tested Negative	Totals
Actually has a tumor	9	1	10
Does not have a tumor	10	80	90
Totals	19	81	100

7. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the

ball cost?

8. In a lake, there is a patch of lilypads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

Appendix E

Frame Conversion Items

Scenario 1: Prescription drug

If an acne cream causes [no] side effects in 8% [92%] of patients, what percent of patients do not [do] experience side effects?

Scenario 2: Academic assessment

If a student answers 28% [72%] of the questions incorrectly [correctly] on a test, what percent of the questions does the student answer correctly [incorrectly]?

Scenario 3: Milk

If a milk is 12% [88%] fat[-free], what percent of the milk is fat-free [fat]?

Scenario 4: Health bars

If 21% [79%] of a health bar does not contain [contains] sugar, what percent of the health bar does [not] contain sugar?